ELSEVIER

Stochastic Processes and their Applications 99 (2002) 117–135

stochastic processes and their applications

www.elsevier.com/locate/spa

Weak convergence for the covariance operators of a Hilbertian linear process

André Mas

CREST-Université Toulouse III, 27, Avenue Bouisson-Bertrand, 34090 Montpellier, France

Received 17 October 2000; received in revised form 8 September 2001; accepted 25 October 2001

Abstract

Let $X_t = \sum_{k=-\infty}^{+\infty} a_k(\varepsilon_{t-k})$ be a linear process with values in a Hilbert space H. The H valued r.v. ε_k are i.i.d. centered, the a_k's are linear operators. We prove a central limit theorem for the vector of empirical covariance operators of the random variables X_t at orders 0 to $h \in \mathbb{N}$ in the space of Hilbert–Schmidt operators. Statistical applications are given in the area of principal component analysis for vector dependent random curves. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Linear operators on Hilbert space; Covariance operators; Weak convergence of random elements

1. Introduction

In the following H will denote a real separable Hilbert space endowed with inner product $\langle , \rangle_H$ and norm $\|.\|_H$. We consider random variables defined on a general probability space $(\Omega, \mathscr{A}, P)$ and with values in H. The setting of our study is the following infinite dimensional stochastic process:

$$X_t = \sum_{k=-\infty}^{+\infty} a_k(\varepsilon_{t-k}). \tag{1}$$

The random variables $(\varepsilon_k)_{k\in\mathbb{Z}}$ are independent, identically distributed and centered. Moment assumptions will be mentioned later. The linear operators a_k map H onto H. We suppose that they are bounded. Note that X_t is a strictly stationary sequence.

This sort of process is of interest when attempting to model continuous time processes. Asymptotic inference when data are time-dependent curves is another statistical application. Since our results clearly fall within the scope of functional data analysis, we

E-mail address: mas@cict.fr (A. Mas).

PII: S0304-4149(02)00087-X

refer to Ramsay and Silverman (1997) for a wide survey about functional techniques in statistics. Consequently in many cases the space H will be a space of functions such as, for instance $L^2([a,b],\sigma([a,b]),\mu)$ where a,b are real numbers ($a<b$), $\sigma([a,b])$ denotes the σ-algebra of Borel sets on $[a,b]$ and μ a finite measure on the real line.

In the special case when the underlying model is autoregressive (namely $X_{t+1}=\rho(X_t)+\varepsilon_{t+1}$ instead of (1)), an exhaustive theoretical study may be found in Bosq (2000) as well as examples.

We will focus on the study of the covariance structure of the process X. Consequently we introduce the following notations: for all u, v in H,

$$u\otimes v=\langle u,.\rangle_H v.$$

The covariance operators under concern are the following

$$C=\mathbb{E}(\varepsilon_0\otimes\varepsilon_0),$$

$$\Gamma_0=\mathbb{E}(X_0\otimes X_0),$$

$$\Gamma_h=\mathbb{E}(X_h\otimes X_0).$$

The linear operators C, Γ and Γ_h are nuclear operators on H when the second order strong moments of the random variables are convergent. The operator

$$\Gamma_{n,h}=\frac{1}{n}\sum_{t=1}^{n}X_{t+h}\otimes X_t$$

is the empirical covariance operator of order h of the process X. We will consider $\Gamma_{n,h}$ as a random variable with values in the space of Hilbert–Schmidt operators. We recall that a bounded operator T on H is Hilbert–Schmidt if and only if the sequence of eigenvalues of the symmetric operator T^*T is summable and that the space of Hilbert–Schmidt operators, denoted $\mathscr{S}$, may be endowed with a (separable) Hilbert space structure. Namely $\langle .,.\rangle$ and $.|\cdot|_2$ will respectively denote the inner product and norm in $\mathscr{S}$. We then have for all T and S in $\mathscr{S}$ and for all complete orthonormal system (u_i) in H:

$$|T|_2^2=\sum_i\|T(u_i)\|_H^2\quad\langle T,S\rangle=\sum_i\langle T(u_i),S(u_i)\rangle_H.$$

We will also make use of the following notation: for all T, S in $\mathscr{S}$,

$$T\tilde{\otimes}S=\langle T,.\rangle S.$$

The operator $T\tilde{\otimes}S$ is nuclear from $\mathscr{S}$ to $\mathscr{S}$. The nuclear norm (resp. the usual norm) of such an operator will be denoted $\|.\|_1$ (resp. $\|.\|_\infty$). We have, for instance $\|T\tilde{\otimes}S\|_1=\|T\|\|S\|$.

Since several operator norms are involved, the reader will just have to remember that norms with single (resp. double) bars will be devoted to operators from H to H (resp. from $\mathscr{S}$ to $\mathscr{S}$). We also refer to the subsection "Elementary facts about operator theory" for basic and common properties of these norms.

Our aim is to prove that the vector of the empirical covariance estimator up to any fixed order

$$\sqrt{n}\begin{pmatrix} \Gamma_{n,0}-\Gamma_0 \\ \Gamma_{n,1}-\Gamma_1 \\ \cdots \\ \Gamma_{n,h}-\Gamma_h \end{pmatrix}$$

converges in distribution (or weakly) in the space $\mathscr{S}^h$ to a Gaussian random element ($\stackrel{w}{\to}$ will denote weak convergence). The case of the empirical mean was treated in Merlevède (1996) and the techniques of proof are quite different.

We suppose from now on that

H.1: $\mathbb{E}\|\varepsilon_0\|_H^4 < +\infty$.

and that

H.2: $\sum_{k=-\infty}^{+\infty} |a_k|_\infty < +\infty$

where, for all bounded linear operator T from H to H, $|T|_\infty$ denotes the operator norm of T.

The particular case when X is a finite dimensional or a real process has been known for a long time. We refer for instance to Brockwell and Davis (1991). Along the proofs some calculations may look like the finite dimensional case, since we will adopt the same techniques (based on m-dependence) as these authors. Nevertheless the functional setting provides many more specificities.

First of all we deal with random operators acting on infinite dimensional spaces. Consequently, all norms are no more equivalent and we have to cope with the unusual metrics on operators spaces mentioned above. Besides, the non-commutative framework provides serious computational problems. At last, the proofs require that the reader is familiar with techniques typically related to weak convergence of measures on infinite dimensional Banach spaces (such as Prohorov's theorem or standard conditions for tightness).

2. Preliminary facts

In this section we recall three "tools" that will be frequently used in proofs. We will denote $\mathscr{H}$ a general real separable Hilbert space (in the sequel we may either have $\mathscr{H}=H$ or $\mathscr{H}=\mathscr{S}$). Notations used are specific to this section. For the sake of simplicity $\mathrm{Var}(Z)$ will denote, for any $\mathscr{H}$ valued random element Z the covariance operator of Z, i.e. $\mathbb{E}(Z\otimes Z)$, and so $\mathrm{Var}(Z)$ will be a linear operator from $\mathscr{H}$ to $\mathscr{H}$.

The first statement in the following theorem is a consequence of a result by Maltsev and Ostrovskii (1982) in the more general setting of stationary mixing sequences in Hilbert space. The second statement may be proved by rewriting the calculations found in Brockwell and Davis (1991).

2.1. The central limit theorem for m-dependent sequences in a Hilbert space

Theorem 1. *Let X_t be a strictly stationary m dependent sequence of random variables with values in $\mathscr{H}$, centered and with autocovariance operator of order $p\Gamma(p)$. If $\Theta_m = \Gamma(0) + \sum_{p=1}^{m}(\Gamma(p) + \Gamma^*(p))$ then, when n goes to infinity*

$$\sqrt{n}\bar{X}_n \xrightarrow{w} G_m, \tag{2}$$

where G_m is a centered Gaussian element in $\mathscr{H}$ with covariance operator Θ_m. We also have

$$nVar(\bar{X}_n) \to \Theta_m \tag{3}$$

in nuclear norm for linear operators on $\mathscr{H}$.

2.2. A theorem of weak convergence for arrays of random variables

The next result is well known and may be found in Billingsley (1968).

Theorem 2. *Let X_{kn} be a doubly indexed sequence of random variables defined on the same probability space as the sequence Y_n. Suppose that all the random variables take values in a separable metric space (S, d) and that*

$$X_{kn} \xrightarrow{w} X_k \quad \text{when } n \to +\infty$$

$$X_k \xrightarrow{w} X \quad \text{when } k \to +\infty$$

$$\lim_k \limsup_n P(d(X_{kn}, Y_n) > \varepsilon) = 0 \quad \text{for all } \varepsilon > 0$$

then

$$Y_n \xrightarrow{w} X.$$

2.3. Elementary facts about operator theory

We deal with linear bounded operators all defined and with values in $\mathscr{H}$ (either H or $\mathscr{S}$). The usual, Hilbert–Schmidt and nuclear norm for linear operators on $\mathscr{H}$ are, respectively denoted $N_\infty(\cdot), N_2(\cdot), N_1(\cdot)$ (representing norms either on H or $\mathscr{S}$). The operators T, S and U (defined and with values in $\mathscr{H}$) are bounded and T is supposed to be either Hilbert–Schmidt or nuclear. The adjoint operator of S is denoted S^*. The following results are elementary in operator theory. They will not be proved.

Fact 1. Let $u, v \in \mathscr{H}$ then if $T = u \otimes v$

$$N_\infty(T) = N_2(T) = N_1(T) = \|u\|_{\mathscr{H}} \|v\|_{\mathscr{H}}. \tag{4}$$

Fact 2. $N_\infty(T) \leqslant N_2(T) \leqslant N_1(T)$ and

$$N_1(UTS) \leqslant N_\infty(U)N_1(T)N_\infty(S),$$

$$N_2(UTS) \leqslant N_\infty(U)N_2(T)N_\infty(S),$$

$$N_\infty(S) = N_\infty(S^*). \tag{5}$$

3. Results

The first result gives a rate of convergence for the cross covariance operator (between $\Gamma_{n,p}$ and $\Gamma_{n,q}$). This cross covariance is an operator from $\mathscr{S}$ to $\mathscr{S}$ and this result is valid in nuclear norm.

Lemma 3. *Let X_t be the following infinite moving average*:

$$X_t = \sum_{j=-\infty}^{+\infty} a_j(\varepsilon_{t-j}) \quad t = 1, \dots, n$$

where the ε_t are i.i.d. centered satisfying H.1 *and the a_j's are bounded operators from H to H satisfying* H.2. *Then*

$$\lim_{n \to +\infty} \| n\mathrm{Cov}(\Gamma_{n,p}, \Gamma_{n,q}) - \Theta_\Gamma^{(p,q)} \|_1 = 0$$

where $\Theta_\Gamma^{(p,q)}$ is defined for all T in $\mathscr{S}$ by

$$\Theta_\Gamma^{(p,q)}(T) = \sum_h \Gamma_{h+p-q} T \Gamma_h + \sum_h \Gamma_{h+q} T \Gamma_{h-p} + A_q(\Lambda - \Phi)A_p(T) \tag{6}$$

where Δ, Φ and A_p are linear operators from $\mathscr{S}$ to $\mathscr{S}$ respectively defined by

$$\Delta(T) = \mathbb{E}((\varepsilon_0 \otimes \varepsilon_0) \tilde{\otimes} (\varepsilon_0 \otimes \varepsilon_0))(T)$$

$$\Phi(T) = C(T + T^*)C + (C \tilde{\otimes} C)(T)$$

and $A_p(T) = \sum_i a_{i+p} T a_i^*$.

Next, we introduce a truncated version of the process X_t. We can apply the CLT for m-dependent sequences to this process. In the next proposition we suppose that $m > h$.

Remark 1. Conversely to the scalar case the random variables under concern (linear operators) do not commute. It follows that it seems impossible to give a more explicit and simpler formulation to the last term, denoted $A_q(\Lambda - \Phi)A_p$.

The next result is an easy consequence of the central limit theorem for m-dependent sequences.

Proposition 4 (Weak convergence for the truncated process). *Let $X_{t,m}$ be the finite order moving average*

$$X_{t,m}=\sum_{j=-m}^{m}a_j(\varepsilon_{t-j}),$$

where the ε_t are i.i.d. centered with $E\|\varepsilon_0\|^4<\infty$ and the a_j are bounded operators from H to H. Let $\Gamma_{n,h,m}$ denote, for any non-negative integer h, the empirical covariance operator of order h of the process $X_{t,m}$ and $\Gamma_{h,m}$ the covariance operator of order h, then

$$\sqrt{n}\begin{pmatrix}\Gamma_{n,0,m}-\Gamma_{0,m}\\ \Gamma_{n,1,m}-\Gamma_{1,m}\\ \cdots\\ \Gamma_{n,h,m}-\Gamma_{h,m}\end{pmatrix}\underset{n\to+\infty}{\overset{w}{\to}}G_{\Gamma,m},$$

where $G_{\Gamma,m}$ is a centered Gaussian element with values in $\mathscr{S}^{h+1}$. The covariance operator of $G_{\Gamma,m}$ is a nuclear operator from $\mathscr{S}^{h+1}$ to $\mathscr{S}^{h+1}$ that may be defined by $(h+1)^2$ blocks (each block is an operator from $\mathscr{S}$ to $\mathscr{S}$) denoted $(\Theta_{\Gamma,m}^{(p,q)})_{0\leqslant p,q\leqslant h}$ where

$$\Theta_{\Gamma,m}^{(p,q)}(T)=\sum_l \Gamma_{l+p-q,m}T\Gamma_{l,m}+\sum_l \Gamma_{l+q,m}T\Gamma_{l-p,m}+A_{q,m}(\Lambda-\Phi)A_{p,m}(T).$$

The operator $A_{p,m}$ is defined by $A_{p,m}=\sum_{-m\leqslant i\leqslant m-p}\theta_{i,i+p}$.

Now we turn to the general case. Theorem 2 is the key to the *main result of this paper*.

Theorem 5 (Weak convergence for the vector of covariances). *Let us consider the following linear and Hilbert space valued process*

$$X_t=\sum_{j=-\infty}^{+\infty}a_j(\varepsilon_{t-j}).$$

Under H.1 *and* H.2

$$\sqrt{n}\begin{pmatrix}\Gamma_{n,0}-\Gamma_{0}\\ \Gamma_{n,1}-\Gamma_{1}\\ \cdots\\ \Gamma_{n,h}-\Gamma_{h}\end{pmatrix}\underset{n\to+\infty}{\overset{w}{\to}}G_{\Gamma}$$

where $G_\Gamma=(G_\Gamma^{(0)},\dots,G_\Gamma^{(h)})$ is a Gaussian centered random element with values in $\mathscr{S}^{h+1}$. Its covariance operator is $\Theta_\Gamma=(\Theta_\Gamma^{(p,q)})_{0\leqslant p,q\leqslant h}$ which is a nuclear operator in $\mathscr{S}^{h+1}$ defined blockwise in Lemma 3.

Remark 2. Note that formula (6) is really similar to formula 7.3.3 in Brockwell and Davis (1991, p. 226) if one sets $\eta\gamma(p)\gamma(q) = A_q \Lambda A_p$ (like Λ, η is the only term related to the fourth moment of ε) and $3\gamma(p)\gamma(q) = A_q \Phi A_p$ (note that Φ is also made of 3 terms directly linked to the covariance of ε). The two other terms (made of explicit infinite sums of covariances) are exactly equivalent.

4. Application to the eigenelements

From the asymptotic normality of the covariance sequence we deduce weak convergence for elements of the principal component analysis of the process X (namely the eigenvalues and the related projection operators of $\Gamma_{n,0} = \Gamma$). We denote λ_i (resp. $\lambda_{i,n}$) the ith eigenvalue of Γ (resp. $\Gamma_n = \Gamma_{n,0}$) arranged in a decreasing order and Π_i (resp. $\Pi_{i,n}$) the associated projector. The proof takes three steps. We refer to Dunford and Schwarz (1988) and to Gohberg et al. (1990) for more information about the theoretical setting surrounding the functional calculus for operators and for defining such terms as "Cauchy contours", "resolvent"... These techniques were already applied in Dauxois et al. (1982) to the principal component analysis of functions in the case of independent random variables.

First step: We note that the inequality

$$\sup_{1 \leqslant i \leqslant n} |\lambda_{i,n} - \lambda_i| \leqslant |\Gamma_{n,0} - \Gamma_0|_\infty,$$

that may be found in Gohberg et al. (1990, p. 99) for instance, enables us to obtain a uniform rate of convergence in probability for the eigenvalues. It is consequently possible to draw disjoint Cauchy contours around each λ_i containing all the $\lambda_{i,n}$ for a sufficiently large n.

Second step: Following Dauxois et al. (1982) we then deduce a CLT for $(\Pi_{i,n} - \Pi_i)$

Third step: Convergence in distribution of $\sqrt{n}(\lambda_{i,n} - \lambda_i)$ is then a corollary of the previous asymptotic results.

Proposition 6 (Weak convergence for the projectors). *We have*

$$\sqrt{n}(\Pi_{i,n} - \Pi_i) \xrightarrow{w} G_\Pi$$

in $\mathscr{S}$ where G_Π is a random operator defined by

$$G_\Pi = S_i G_\Gamma \Pi_i + \Pi_i G_\Gamma S_i$$

where S_i is a continuous linear application on H defined by

$$S_i = \sum_{l \neq i} (\lambda_i - \lambda_l)^{-1} \Pi_l.$$

Proposition 7 (Weak convergence for the eigenvalues). *Suppose that λ_i is an eigenvalue of order one. Then,*

$$\sqrt{n}(\lambda_{i,n} - \lambda_i) \xrightarrow{w} N_i.$$

where N_i is a real, Gaussian, centered random variable with variance

$$\sigma_i^2 = \mathbb{E}\langle G_\Gamma^{(0)}(e_p), e_p\rangle^2.$$

5. Proofs

Proof of Lemma 3. We have

$$\Gamma_{n,p}\tilde{\otimes}\Gamma_{n,q} = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n} K_{s,t}(p,q),$$

where

$$\begin{aligned} &K_{s,t}(p,q)\\ &\quad= \left[\sum_{i=-\infty}^{+\infty} a_i(\varepsilon_{t-i}) \otimes \sum_{j=-\infty}^{+\infty} a_j(\varepsilon_{t+p-j})\tilde{\otimes}\sum_{k=-\infty}^{+\infty} a_k(\varepsilon_{s-k}) \otimes \sum_{l=-\infty}^{+\infty} a_l(\varepsilon_{s+q-l})\right]\\ &\quad=\sum_{i,j}\theta_{i,j}(\varepsilon_{t-i}\otimes\varepsilon_{t+p-j})\tilde{\otimes}\sum_{k,l}\theta_{k,l}(\varepsilon_{s-k}\otimes\varepsilon_{s+q-l}) \end{aligned}$$

and where $\theta_{i,j}$ is a linear mapping from $\mathscr{S}$ to $\mathscr{S}$ defined by

$$\theta_{i,j}(T) = a_j T a_i^*.$$

Note that, with our notations and by stationarity, for all p, h

$$\Gamma_p = \sum_{j\in\mathbb{Z}}\theta_{j,p+j}(C). \tag{7}$$

Finally

$$\begin{aligned} K_{s,t}(p,q) &= \sum_{i,j,k,l}[\theta_{i,j}(\varepsilon_{t-i}\otimes\varepsilon_{t+p-j})]\tilde{\otimes}[\theta_{k,l}(\varepsilon_{s-k}\otimes\varepsilon_{s+q-l})]\\ &= \sum_{i,j,k,l}\theta_{s-k,s+q-l}[(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_k\otimes\varepsilon_l)]\theta^*_{t-i,t+p-j}. \end{aligned}$$

We denote

$$I_1 = \{(i,j,k,l)\colon i=j=k=l\},$$

$$I_2 = \{(i,j,k,l)\colon i=j\neq k=l\},$$

$$I_3 = \{(i,j,k,l)\colon i=k\neq j=l\},$$

$$I_4 = \{(i,j,k,l)\colon i=l\neq j=k\}.$$

Note that

$$\mathbb{E}((\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_k\otimes\varepsilon_l)) = 0$$

when

$$(i,j,k,l) \notin I_1 \cup I_2 \cup I_3 \cup I_4.$$

Now, when $i=j=k=l$,

$$\mathbb{E}((\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_k \otimes \varepsilon_l)) = \mathbb{E}((\varepsilon_0 \otimes \varepsilon_0)\tilde{\otimes}(\varepsilon_0 \otimes \varepsilon_0)) = \Lambda.$$

We also have (in the case corresponding to I_2):

$$\mathbb{E}((\varepsilon_i \otimes \varepsilon_i)\tilde{\otimes}(\varepsilon_k \otimes \varepsilon_k)) = C\tilde{\otimes}C$$

when $i \neq k$. We have to determine (case I_3):

$$\mathbb{E}((\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)).$$

Denote (e_i) a basis of eigenvectors of C. Take an operator $T = \sum_{m,r} t_{m,r} e_m \otimes e_r$ in $\mathscr{S}$

$$\begin{aligned}
&\mathbb{E}((\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j))(T) \\
&\quad = \mathbb{E}\left(\sum_l (\langle \varepsilon_i, e_l\rangle_H \langle T(e_l), \varepsilon_j\rangle_H)(\varepsilon_i \otimes \varepsilon_j)\right) \\
&\quad = \mathbb{E}\left(\sum_{l,r} (t_{l,r}\langle \varepsilon_i, e_l\rangle_H \langle e_r, \varepsilon_j\rangle_H)(\varepsilon_i \otimes \varepsilon_j)\right) \\
&\quad = \sum_{l,r} (t_{l,r}\mathbb{E}(\langle \varepsilon_i, e_l\rangle_H \langle \varepsilon_i, .\rangle)\mathbb{E}(\langle e_r, \varepsilon_j\rangle_H \varepsilon_j)) \\
&\quad = \sum_{l,r} \lambda_l \lambda_r t_{l,r}\langle e_l, .\rangle_H e_r \\
&\quad = CTC.
\end{aligned}$$

Similar calculations lead to

$$\mathbb{E}((\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_j \otimes \varepsilon_i))(T) = CT^*C.$$

We get

$$\begin{aligned}
&\mathbb{E}K_{s,t}(p,q) \\
&\quad = \sum_{(i,j,k,l)\in I_1\cup I_2\cup I_3\cup I_4} \theta_{s-k,s+q-l}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_k \otimes \varepsilon_l)]\theta^*_{t-i,t+p-j}.
\end{aligned}$$

But

$$\begin{aligned}
&\sum_{(i,j,k,l)\in I_2} \theta_{s-k,s+q-l}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_k \otimes \varepsilon_l)]\theta^*_{t-i,t+p-j} \\
&\quad = \sum_i \theta_{t-i,t+p-i}\mathbb{E}(\varepsilon_i \otimes \varepsilon_i)\tilde{\otimes}\sum_k \theta_{s-k,s+q-k}\mathbb{E}(\varepsilon_k \otimes \varepsilon_k)
\end{aligned}$$

$$-\sum_i (\theta_{t-i,t+p-i}C)\tilde{\otimes}(\theta_{s-i,s+q-i}C)$$

$$= \Gamma_p \tilde{\otimes} \Gamma_q - \sum_i \theta_{s-i,s+q-i}(C\tilde{\otimes}C)\theta^*_{t-i,t+p-i}$$

by (7).

Note that

$$\mathbb{E}(\Gamma_{n,p})\tilde{\otimes}\mathbb{E}(\Gamma_{n,q}) = \Gamma_p \tilde{\otimes} \Gamma_q.$$

Now

$$\begin{aligned}
\mathrm{Cov}(\Gamma_{n,p},\Gamma_{n,q}) &= \mathbb{E}(\Gamma_{n,p}\tilde{\otimes}\Gamma_{n,q}) - \mathbb{E}(\Gamma_{n,p})\tilde{\otimes}\mathbb{E}(\Gamma_{n,q}) \\
&= \sum_{(i,j,k,l)\in I_1\cup I_3\cup I_4} \theta_{s-k,s+q-l}\mathbb{E}[(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_k\otimes\varepsilon_l)]\theta^*_{t-i,t+p-j} \\
&\quad -\sum_i \theta_{s-i,s+q-i}(C\tilde{\otimes}C)\theta^*_{t-i,t+p-i} \\
&= \sum_{i\in I_1} \theta_{s-i,s+q-i}\mathbb{E}[(\varepsilon_i\otimes\varepsilon_i)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_i)]\theta^*_{t-i,t+p-i} \\
&\quad -\sum_i \theta_{s-i,s+q-i}(C\tilde{\otimes}C)\theta^*_{t-i,t+p-i} \\
&\quad +\sum_{i\neq j} \theta_{s-i,s+q-j}\mathbb{E}[(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_j)]\theta^*_{t-i,t+p-j} \\
&\quad +\sum_{i\neq j} \theta_{s-j,s+q-i}\mathbb{E}[(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_j\otimes\varepsilon_i)]\theta^*_{t-i,t+p-j}.
\end{aligned}$$

We are going to compute the last two terms first.

Take an operator T in $\mathscr{S}, \theta^*_{t-i,p+t-j}(T) = a^*_{p+t-j}Ta_{t-i}$ then

$$[\mathbb{E}((\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_j))\theta^*_{t-i,t+p-j}](T) = Ca^*_{p+t-j}Ta_{t-i}C$$

and

$$\begin{aligned}
&[\theta_{s-i,s+q-j}\mathbb{E}((\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_j))\theta^*_{t-i,t+p-j}](T) \\
&\quad = a_{s+q-j}Ca^*_{p+t-j}Ta_{t-i}Ca^*_{s-i},
\end{aligned}$$

hence

$$\begin{aligned}
&\sum_{i\neq j} \theta_{s-i,s+q-j}\mathbb{E}[(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_j)]\theta^*_{t-i,t+p-j}(T) \\
&\quad = \left(\sum_j a_{s+q-j}Ca^*_{p+t-j}\right) T \left(\sum_i a_{t-i}Ca^*_{s-i}\right)
\end{aligned}$$

$$-\sum_i \theta_{s-i,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)]\theta^*_{t-i,t+p-i}(T)$$

$$=\Gamma_{s-t+q-p}T\Gamma_{s-t} - \sum_i \theta_{s-i,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)]\theta^*_{t-i,t+p-i}(T).$$

Similar calculations would lead to

$$\theta_{s-j,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_j \otimes \varepsilon_i)]\theta^*_{t-i,t+p-j}(T)$$
$$=a_{s+q-i}Ca^*_{t-i}T^* a_{s+q-j}Ca^*_{s-j}$$

so

$$\sum_{i\neq j} \theta_{s-j,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_j \otimes \varepsilon_i)]\theta^*_{t-i,t+p-j}(T)$$
$$=\Gamma_{s-t+q}T^*\Gamma_{s-t-p} - \sum_i \theta_{s-i,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_j \otimes \varepsilon_i)]\theta^*_{t-i,t+p-i}(T).$$

The last term to be computed is

$$\sum_i \theta_{s-i,s+q-i}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_i)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_i)]\theta^*_{t-i,t+p-i}$$
$$=\sum_i \theta_{s-t+i,s-t+q-i}\mathbb{E}[(\varepsilon_0 \otimes \varepsilon_0)\tilde{\otimes}(\varepsilon_0 \otimes \varepsilon_0)]\theta^*_{i,p+i}$$
$$=\sum_i \theta_{s-t+i,s-t+q-i}\Lambda\theta^*_{i,p+i}.$$

Collecting our three last results we get

$$n\mathrm{Cov}(\Gamma_{n,p},\Gamma_{n,q})(T)$$
$$=\frac{1}{n}\sum_{l=0}^{n-1}(n-l)(\Gamma_{l+p-q}T\Gamma_l + \Gamma_{-l+p-q}T\Gamma_{-l})$$
$$+\frac{1}{n}\sum_{l=0}^{n-1}(n-l)(\Gamma_{l+q}T\Gamma_{l-p} + \Gamma_{-l+q}T\Gamma_{-l-p})$$
$$+\left[\sum_i \frac{1}{n}\sum_{l=-(n-1)}^{(n-1)}(n-l)\theta_{i+l,i+l+q}(\Lambda-\Phi)\theta^*_{i,i+p}\right](T)$$

which may be rewritten

$$n\mathrm{Cov}(\Gamma_{n,p},\Gamma_{n,q})$$
$$=\frac{1}{n}\sum_{l=0}^{n-1}(n-l)[\Psi_{l+p-q,l}+\Psi_{-l+p-q,-l}+\Psi_{l+q,l-p}+\Psi_{-l+q,-l-p}]$$
$$+\left[\sum_i \frac{1}{n}\sum_{l=-(n-1)}^{(n-1)}(n-l)\theta_{i+l,i+l+q}(\Lambda-\Phi)\theta^*_{i,i+p}\right] \tag{8}$$

where

$$\Psi_{l,l'}(T)=\Gamma_l T\Gamma_{l'}.$$

We need two lemmas in order to finish the proof.

Lemma 8. *The following hold for* $0\leqslant p,q\leqslant h$

$$\sum_{l\in\mathbb{Z}}(\|\Psi_{l+p-q,l}\|_1+\|\Psi_{l+q,l-p}\|_1)<+\infty,$$
$$\sum_{l\in\mathbb{Z}}\|\theta_{l,l+q}(\Lambda-\Phi)\|_1<+\infty. \tag{9}$$

Lemma 9. *Equivalent results for the truncated sequence are*

$$\sum_{l\in\mathbb{Z}}(\|\Psi_{l+p-q,l,m}\|_1+\|\Psi_{l+q,l-p,m}\|_1)<+\infty,$$
$$\sum_{l\in\mathbb{Z}}\|\theta_{l,l+q,m}(\Lambda-\Phi)\|_1<+\infty.$$

For obvious reasons, we will restrict ourselves to the proof of Lemma 8.

Proof of Lemma 8. Let us first deal with the second term first:

$$\|\theta_{l,l+p}(\Lambda-\Phi)\|_1\leqslant\|\Lambda-\Phi\|_1\|\theta_{l,l+p}\|_\infty$$
$$\leqslant(\|\Lambda\|_1+\|\Phi\|_1)|a_l|_\infty|a_{l+p}|_\infty.$$

The first inequality is due to (5) and

$$\|\Lambda\|_1\leqslant\mathbb{E}\|(\varepsilon_i\otimes\varepsilon_j)\tilde{\otimes}(\varepsilon_i\otimes\varepsilon_j)\|_1$$
$$=\mathbb{E}\|\varepsilon_0\|_H^4.$$

Elementary calculations lead to

$$\|\Phi\|_1\leqslant 3(\mathbb{E}\|\varepsilon_0\|_H^2)^2.$$

By (4) the inequality

$$\|\theta_{i,j}\|_\infty \leqslant |a_i|_\infty |a_j|_\infty$$

is valid and

$$\sum_{l\in\mathbb{Z}} \|\theta_{l,l+q}\|_1 < +\infty$$

since $\sum |a_l|_\infty < +\infty$.

We now prove that the first series in (9) converges. Note that we proved above that

$$\Psi_{l+p-q,l} = \sum_{i,j} \theta_{l-i,l+q-j}\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)]\theta^*_{-i,p-j}$$

hence

$$\begin{aligned}\sum_l \|\Psi_{l+p-q,l}\|_1 &\leqslant \sum_l \sum_{i,j} \|\theta_{l-i,l+q-j}\|_\infty \|\theta^*_{-i,p-j}\|_\infty \|\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)]\|_1 \\ &\leqslant (\mathbb{E}\|\varepsilon_0\|^2)^2 \sum_{i,j,l} |a_{l-i}|_\infty |a_{l+q-j}|_\infty |a_{-i}|_\infty |a_{p-j}|_\infty,\end{aligned}$$

where once again we made use of (5) in the first inequality and of (4) in the second since:

$$\begin{aligned}\|\mathbb{E}[(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)]\|_1 &\leqslant \mathbb{E}\|(\varepsilon_i \otimes \varepsilon_j)\tilde{\otimes}(\varepsilon_i \otimes \varepsilon_j)\|_1 \\ &= \mathbb{E}|\varepsilon_i \otimes \varepsilon_j|_2^2 \\ &= \mathbb{E}(\|\varepsilon_i\|_H^2 \|\varepsilon_j\|_H^2).\end{aligned}$$

The infinite triple sum is bounded by

$$(\sup \|a_k\|_\infty) \sum_j |a_j|_\infty \sum_{i,l} |a_{l-i}|_\infty |a_{-i}|_\infty < +\infty$$

which implies

$$\sum_{l\in\mathbb{Z}} \|\Psi_{l+p-q,l}\|_1 < +\infty.$$

It is clearly possible to deal exactly the same way with the second term in the sum (9) to obtain

$$\sum_{l\in\mathbb{Z}} \|\Psi_{l+q,l-p}\|_1 < +\infty,$$

which concludes the proof of Lemma 8. □

We just need to invoke Lebesgue's dominated convergence together with Lemmas 8 and 9 to conclude that the limit in (8) is $\Theta_\Gamma^{(p,q)}$ for the nuclear norm of operators on $\mathscr{S}$ which finishes the proof of Proposition 3. □

Proof of Proposition 4. The space $\mathscr{S}^{h+1}$ may be viewed as a—somewhat complicated—usual separable Hilbert space. The process X_t is a $2m+1$-dependent sequence in H. It is also clear that for all $k = 0, \dots, h$, $X_t \otimes X_{t+k}$ is a $2m + h + 1$ dependent sequence in $\mathscr{S}$. Consequently we are under the hypotheses of Proposition 1. The covariance operator of $G_{\Gamma,m}$ must be computed now. Calculations are exactly the same as those of Lemma 3. We just have to note that $\Gamma_{l,m} = 0$ for $|l| > 2m$. □

Proof of Theorem 5. We denote

$$X_{t,m} = \sum_{j=-m}^{m} a_j(\varepsilon_{t-j})$$

and

$$\Gamma_{n,p,m} = \frac{1}{n}\sum_{t=1}^{n} X_{t,m} X_{t+p,m}.$$

Then by Proposition 4

$$\sqrt{n}\begin{pmatrix} \Gamma_{n,0,m} - \Gamma_{0,m} \\ \Gamma_{n,1,m} - \Gamma_{1,m} \\ \dots \\ \Gamma_{n,h,m} - \Gamma_{h,m} \end{pmatrix} \underset{n\to+\infty}{\overset{w}{\to}} G_{\Gamma,m}$$

where $G_{\Gamma,m}$ is centered and Gaussian. Its covariance operator is $\Theta_{\Gamma,m}$. Finally, if we prove the two following convergence results:

$$G_{\Gamma,m} \overset{w}{\to} G_\Gamma \quad \text{in } \mathscr{S}^h, \tag{10}$$

when m tends to infinity and

$$\lim_{m\to+\infty} \limsup_{n\to+\infty} P(\sqrt{n}|\Gamma_{n,p,m} - \Gamma_{p,m} - \Gamma_{n,p} + \Gamma_p|_2 > \varepsilon) = 0 \tag{11}$$

for all p and ε, it will then be sufficient to invoke Theorem 2 to conclude. First dealing with (11) we get

$$\begin{aligned} &P(\sqrt{n}|\Gamma_{n,p,m} - \Gamma_{p,m} - \Gamma_{n,p} + \Gamma_p|_2 > \varepsilon) \\ &\quad\leqslant \frac{n}{\varepsilon^2}\mathbb{E}|\Gamma_{n,p,m} - \Gamma_{p,m} - \Gamma_{n,p} + \Gamma_p|_2^2 \\ &\quad= \frac{n}{\varepsilon^2}[\mathbb{E}|\Gamma_{n,p,m} - \Gamma_{p,m}|_2^2 + \mathbb{E}|\Gamma_{n,p} - \Gamma_p|_2^2 - 2\mathbb{E}\langle \Gamma_{n,p,m} - \Gamma_{p,m}, \Gamma_{n,p} - \Gamma_p\rangle]. \end{aligned}$$

Lemma 8 yields

$$n\mathbb{E}|\Gamma_{n,p} - \Gamma_p|_2^2 = \|n\mathrm{Cov}(\Gamma_{n,p}, \Gamma_{n,p})\|_1 \to \|\Theta_\Gamma^{(p,p)}\|_1$$

and Lemma 9

$$n\mathbb{E}|\Gamma_{n,p,m} - \Gamma_{p,m}|_2^2 \to \|\Theta_{\Gamma,m}^{(p,p)}\|_1.$$

Besides, it is plain that

$$\mathbb{E}\langle \Gamma_{n,p,m} - \Gamma_{p,m}, \Gamma_{n,p} - \Gamma_p\rangle = \mathbb{E}|\Gamma_{n,p,m} - \Gamma_{p,m}|_2^2.$$

So if we prove the following lemma (which will be done below) we will obtain (11) as a by-product.

Lemma 10. *For all* $p \in \{0, \ldots, h\}$

$$\lim_{m \to +\infty} \|\Theta_{\Gamma,m}^{(p,p)} - \Theta_{\Gamma}^{(p,p)}\|_1 = 0.$$

Now turning to (10), Lemma 10 implies clearly weak convergence for all the finite dimensional distributions. We just have to prove that $G_{\Gamma,m}$ is a weakly relatively compact sequence or equivalently here that it is flatly concentrated.

We recall the a sequence μ_n of probability measures on a Banach space B is flatly concentrated if, for all $\varepsilon > 0$, there exists a finite dimensional vector space $B_0(\varepsilon)$ such that

$$\inf_n \mu_n(B_0^{\varepsilon}(\varepsilon)) > 1 - \varepsilon$$

where B_0^{ε} denotes the ε-neighborhood of B_0.

It then suffices to prove that all the coordinates of the random vector $G_{\Gamma,m}$ are flatly concentrated. Once more Lemma 10 will be the key. Denote $G_{\Gamma,m}^{(p)} \in \mathscr{S}$ (resp. $G_{\Gamma}^{(p)}$) the pth (of $h+1$) coordinate of $G_{\Gamma,m}$ (resp. of G_{Γ}) and $\mathscr{P}_k^{(p)}$ the projection operator on the k first eigenvector of the covariance operator of $G_{\Gamma}^{(p)}$. We claim that

$$\lim_k \limsup_m P(|(I - \mathscr{P}_k^{(p)}) G_{\Gamma,m}^{(p)}|_2 > \varepsilon) = 0. \tag{12}$$

Indeed

$$\begin{aligned}
&P(|(I - \mathscr{P}_k^{(p)}) G_{\Gamma,m}^{(p)}|_2 > \varepsilon)\\
&\quad \leqslant \mathbb{E}|(I - \mathscr{P}_k^{(p)}) G_{\Gamma,m}^{(p)}|_2^2\\
&\quad = \|(I - \mathscr{P}_k^{(p)}) \Theta_{\Gamma,m}^{(p,p)} (I - \mathscr{P}_k^{(p)})\|_1\\
&\quad \leqslant \|(I - \mathscr{P}_k^{(p)})[\Theta_{\Gamma,m}^{(p,p)} - \Theta_{\Gamma}^{(p,p)}](I - \mathscr{P}_k^{(p)})\|_1 + \|(I - \mathscr{P}_k^{(p)}) \Theta_{\Gamma}^{(p,p)} (I - \mathscr{P}_k^{(p)})\|_1\\
&\quad \leqslant \|\Theta_{\Gamma,m}^{(p,p)} - \Theta_{\Gamma}^{(p,p)}\|_1 + \|(I - \mathscr{P}_k^{(p)}) \Theta_{\Gamma}^{(p,p)} (I - \mathscr{P}_k^{(p)})\|_1.
\end{aligned}$$

By Lemma 10 the first term tends to zero without depending on k and the second also tends to zero as k tend to infinity. This proves (12) that ensures that the sequence of Gaussian random variables $(G_{\Gamma,m}^{(p)})_m$ is flatly concentrated. We obtain (10).

Our last task consists in proving Lemma 10.

Proof of Lemma 10.

$$\Theta_{\Gamma,m}^{(p,p)} = \sum_l \Psi_{l,l,m} + \sum_l \Psi_{l+p,l-p,m} + A_{p,m}(\Lambda - \Phi)A_{p,m}$$

and

$$\begin{aligned}
\Theta_{\Gamma,m}^{(p,p)} - \Theta_{\Gamma}^{(p,p)} = {} & \sum_l (\Psi_{l,l,m} - \Psi_{l,l}) + \sum_l (\Psi_{l+p,l-p,m} - \Psi_{l+p,l-p})\\
& + (A_{p,m} - A_p)(\Lambda - \Phi)A_{p,m} - A_p(\Lambda - \Phi)(A_p - A_{p,m}).
\end{aligned}$$

We have

$$\|(A_{p,m} - A_p)\Lambda A_{p,m}\|_1 \leqslant \|\Lambda - \Phi\|_1 \|A_{p,m} - A_p\|_\infty \|A_{p,m}\|_\infty$$

$$A_{p,m} - A_p = \sum_{i<-m,\ i>m-p} \theta_{i,i+p}$$

and clearly

$$\|A_{p,m} - A_p\|_\infty \leqslant \sum_{i<-m,\ i>m-p} |a_i|_\infty |a_{i+p}|_\infty$$

$$\leqslant (\sup |a_k|_\infty) \sum_{i<-m,\ i>m} |a_i|_\infty$$

and

$$\lim_{m\to+\infty} \|A_{p,m} - A_p\|_\infty = \lim_{m\to+\infty} \|(A_{p,m} - A_p)(\Lambda - \Phi)A_{p,m}\|_1$$

$$= \lim_{m\to+\infty} \|A_p(\Lambda - \Phi)(A_p - A_{p,m})\|_1 = 0.$$

Let us turn to the first term

$$(\Psi_{l,l,m} - \Psi_{l,l})(T) = \Gamma_l T \Gamma_l - \Gamma_{l,m} T \Gamma_{l,m}$$

$$= (\Gamma_l - \Gamma_{l,m}) T \Gamma_l + \Gamma_{l,m} T (\Gamma_l - \Gamma_{l,m})$$

and prove

$$\lim_{m\to+\infty} \sum_l \|\tilde{\Psi}_{l,l,m}\|_1 = 0$$

where $\tilde{\Psi}_{l,l,m}(T) = (\Gamma_l - \Gamma_{l,m}) T \Gamma_l$ and

$$\Gamma_l - \Gamma_{l,m} = \begin{cases} \displaystyle\sum_{i<-m,\ i>m-l} \theta_{i,i+l} C & l \geqslant 0, \\ \displaystyle\sum_{i<-m-l,\ i>m} \theta_{i,i+l} C & l < 0. \end{cases}$$

So

$$(\Gamma_l - \Gamma_{l,m}) T \Gamma_l = \begin{cases} \displaystyle\sum_{i<-m,\ i>m-l} \sum_j a_i C a_{i+l} T a_j C a_{j+l} & l \geqslant 0, \\ \displaystyle\sum_{i<-m-l,\ i>m} \sum_j a_i C a_{i+l} T a_j C a_{j+l} & l < 0. \end{cases}$$

We denote κ and $\kappa_{i,j,l}$ the linear mappings from $\mathscr{S}$ to $\mathscr{S}$ defined, respectively, by $\kappa(T) = CTC$ and $\kappa_{i,j,l}(T) = a_i C a_{i+l} T a_j C a_{j+l}$ we have

$$\kappa_{i,j,l} = \theta_{i,j+l} \kappa \theta_{i+l,j}$$

$$\|\kappa_{i,j,l}\|_1 \leqslant \|\theta_{i,j+l}\|_\infty \|\kappa\|_1 \|\theta_{i+l,j}\|_\infty$$

$$\leqslant \|\kappa\|_1 |a_i|_\infty |a_{i+l}|_\infty |a_j|_\infty |a_{j+l}|_\infty$$

where we already proved that $\|\kappa\|_1 \leqslant (\mathbb{E}\|\varepsilon_0\|^2)^2$. Hence

$$\sum_l \|\tilde{\Psi}_{l,l,m}\|_1 \leqslant \|\kappa\|_1 (\sup \|a_k\|_\infty) \left(\sum_j |a_j|_\infty\right)$$
$$\times \left(\sum_{l \geqslant 0} \sum_{i<-m,\ i>m-l} |a_i|_\infty |a_{i+l}|_\infty + \sum_{l<0} \sum_{i<-m-l,\ i>m} |a_i|_\infty |a_{i+l}|_\infty\right)$$

and clearly

$$\left(\sum_{l \geqslant 0} \sum_{i<-m,\ i>m-l} |a_i|_\infty |a_{i+l}|_\infty + \sum_{l<0} \sum_{i<-m-l,\ i>m} |a_i|_\infty |a_{i+l}|_\infty\right)$$
$$\leqslant 2\left(\sum_l |a_l|_\infty\right)\left(\sum_{i<-m} |a_i|_\infty + \sum_{i>m} |a_i|_\infty\right)$$

that yields

$$\lim_{m \to +\infty} \sum_l \|\tilde{\Psi}_{l,l,m}\|_1 = 0$$

and finishes the proof of Lemma 10. □

Proof of Proposition 6. We denote $\Gamma_0 = \Gamma$ and $\Gamma_{n,0} = \Gamma_n$. We denote $\mathscr{C}_j$ (resp. $\mathscr{C}_{j,n}$) a deterministic (resp. random) Cauchy contour around λ_j (resp. $\lambda_{j,n}$) such that $\mathscr{C}_j$ (resp. $\mathscr{C}_{j,n}$) contains no other eigenvalue then

$$\Pi_{j,n} = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}_{j,n}} (zI - \Gamma_n)^{-1}\, \mathrm{d}z,$$
$$\Pi_j = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}_j} (zI - \Gamma)^{-1}\, \mathrm{d}z.$$

Suppose furthermore that

$$\Pi_{j,n} = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}_j} (zI - \Gamma_n)^{-1}\, \mathrm{d}z \tag{13}$$

then

$$\Pi_{j,n} - \Pi_j = \frac{1}{2\pi \mathrm{i}} \int_{\mathscr{C}_j} (zI - \Gamma)^{-1} (\Gamma_n - \Gamma)(zI - \Gamma)^{-1}\, \mathrm{d}z + R_n,$$

where

$$R_n = \frac{1}{2\pi \mathrm{i}} \left(\int_{\Gamma_j} \{(\lambda - \Gamma)^{-1} n^{1/4} (\Gamma - \Gamma_n)(\lambda - \Gamma)^{-1} n^{1/4} (\Gamma - \Gamma_n)(\lambda - \Gamma_n)^{-1}\}\, \mathrm{d}\lambda\right)$$

and consequently

$$\|R_n\|_{\mathscr{S}} = \mathrm{O}(\|\Gamma_n - \Gamma\|_{\mathscr{S}}^2).$$

Let us introduce the following bounded linear mapping from $\mathscr{S}$ to $\mathscr{S}\times\mathscr{S}$

$$\varphi_j(T)=\begin{pmatrix}\dfrac{1}{2\pi\mathrm{i}}\displaystyle\int_{\mathscr{C}_j}(zI-\Gamma)^{-1}T(zI-\Gamma)^{-1}\,\mathrm{d}z\\ T\end{pmatrix}.$$

P-continuity theorems imply that for any fixed j,

$$\varphi_j(\sqrt{n}(\Gamma_n-\Gamma))=\sqrt{n}\begin{pmatrix}\Pi_{j,n}-\Pi_j\\ \Gamma_n-\Gamma\end{pmatrix}$$

converges weakly to the Gaussian random element $\varphi_j(G_\Gamma)$. So

$$\sqrt{n}(\Pi_{j,n}-\Pi_j)\to_w \frac{1}{2\pi\mathrm{i}}\int_{\mathscr{C}_j}(zI-\Gamma)^{-1}G_\Gamma(zI-\Gamma)^{-1}\,\mathrm{d}z.$$

Calculations made in Dauxois et al. (1982, p. 145) give the announced result. Note that the event

$$\left\{\Pi_{j,n}=\frac{1}{2\pi\mathrm{i}}\int_{\mathscr{C}_j}(zI-\Gamma_n)^{-1}\,\mathrm{d}z\right\}$$

is also the event $\Omega_n=\{\mathscr{C}_{i,n}\text{ is inside }\mathscr{C}_i\}$.

In order to conclude we just need to prove that

$$\varphi_j(\sqrt{n}(\Gamma_n-\Gamma)\mathbb{1}_{\Omega_n^c})\to_P 0,$$

where $\mathbb{1}_O$ denotes the indicator function of the set O. But

$$\mathbb{E}\|\varphi_j(\sqrt{n}(\Gamma_n-\Gamma)\mathbb{1}_{\Omega_n^c})\|\leqslant\|\varphi_j\|_\infty\sqrt{\mathbb{E}\|\sqrt{n}(\Gamma_n-\Gamma)\|}\sqrt{P(\Omega_n^c)}.$$

Consequently it suffices to prove that

$$P\left(\Pi_{j,n}=\frac{1}{2\pi\mathrm{i}}\int_{\mathscr{C}_j}(zI-\Gamma_n)^{-1}\,\mathrm{d}z\right)\to 1,$$

as n tends to infinity. In fact, for the sake of simplicity we take for $\mathscr{C}_{j,n}$ the circle with center $\lambda_{j,n}$ and radius $\rho_j=\frac{1}{4}\min(\lambda_j-\lambda_{j+1},\lambda_{j-1}-\lambda_j)$ and for $\mathscr{C}_j$ the circle with center λ_j and radius $2\rho_j$. Now since $|\lambda_j-\lambda_{j,n}|=\mathrm{O}(1/\sqrt{n})$ we may suppose that for a sufficiently large $n|\lambda_j-\lambda_{j,n}|<\rho_j/2$. A rough sketch finishes the proof. □

Proof of Proposition 7. We refer once more to Dauxois et al. (1982, p. 147–148). Here the proof is quite similar and omitted. □

6. Uncited References

Mas, 1999; Rudin, 1987; Ruymgaart and Yang, 1997; Skorohod, 1984.

Acknowledgements

I am grateful to the anonymous referee for helpful comments that made the article gain in clarity.

References

Billingsley, P., 1968. Convergence of Probability Measures. Wiley Series in Probability and Mathematical Statistics. Wiley, New York.

Bosq, D., 2000. Linear Processes in Function spaces. Lectures notes in Statistics. Springer, Berlin.

Brockwell, P.J., Davis, R.A., 1991. Times Series: Theory and Method. Springer Series in Statistics. Springer, Berlin.

Dauxois, J., Pousse, A., Romain, Y., 1982. Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. J. Multivar. Anal. 12, 136–154.

Dunford, N., Schwarz, J.T., 1988. Linear operators, Vols. I–III. Wiley Classics Library. Wiley, New York.

Gohberg, I., Goldberg, S., Kaashoek, M.A., 1990. Classes of Linear Operators, Vol I. Operator Theory: Advances and Applications. Vol 49. Birkhaüser Verlag, New York.

Maltsev, V.V., Ostrovskii, E.I., 1982. Central limit theorem for stationary processes in Hilbert space. Theor. Probab. Appl. 27, 357–359.

Mas, A., 1999. Principal component analysis in Hilbert space: a perturbation approach. Document de Travail CREST No. 9970.

Merlevède, F., 1996. Central limit theorem for linear processes with values in a Hilbert space. Stochastic Processes Appl. 65 (1), 103–114.

Ramsay, J.O., Silverman, B.W., 1997. Functional Data Analysis. Springer Series in Statistics, Berlin.

Rudin, W., 1987. Complex and real Analysis, 3rd Edition. McGraw-Hill, New York.

Ruymgaart, F.H., Yang, S., 1997. Some applications of Watson's perturbation approach to random matrices. J. Multivar. Anal. 60, 48–60.

Skorohod, A.N, 1984. Random Linear Operators. Mathematics and its Applications. D. Reidel Publishing.